

## Research Article

# Dynamical Criticality in Gene Regulatory Networks

**Marco Villani** <sup>1,2</sup>, **Luca La Rocca**<sup>1</sup>, **Stuart Alan Kauffman**<sup>3</sup>, and **Roberto Serra**<sup>1,2</sup>

<sup>1</sup>Department of Physics, Informatics and Mathematics, University of Modena and Reggio Emilia, via Campi 213a, 41125 Modena, Italy

<sup>2</sup>European Centre for Living Technology, San Marco 2940, 30124 Venezia, Italy

<sup>3</sup>Institute for Systems Biology, Seattle 401 Terry Ave N, Seattle, WA 98109, USA

Correspondence should be addressed to Marco Villani; marco.villani@unimore.it

Received 22 April 2018; Revised 3 August 2018; Accepted 27 August 2018; Published 4 October 2018

Academic Editor: Eulalia Martínez

Copyright © 2018 Marco Villani et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

A well-known hypothesis, with far-reaching implications, is that biological evolution should preferentially lead to states that are dynamically critical. In previous papers, we showed that a well-known model of genetic regulatory networks, namely, that of random Boolean networks, allows one to study in depth the relationship between the dynamical regime of a living being's gene network and its response to permanent perturbations. In this paper, we analyze a huge set of new experimental data on single gene knockouts in *S. cerevisiae*, laying down a statistical framework to determine its dynamical regime. We find that the *S. cerevisiae* network appears to be slightly ordered, but close to the critical region. Since our analysis relies on dichotomizing continuous data, we carefully consider the issue of an optimal threshold choice.

## 1. Introduction

The idea that dynamically critical states possess some peculiar advantages, so that they tend to be selected under evolutionary dynamics, has relevant implications in biology [1, 2] as well as in evolutionary computation and in the design of artificial systems [3]. This idea, which will be referred to as the *criticality hypothesis* in this paper, has been suggested by several authors [1–6] and has played a prominent role in the search for general principles in complexity science [7–9]. It is based on the observation that, in nonlinear dynamical systems, the same set of equations can lead to very different behaviours (depending upon the values of some parameters).

Let us consider in particular dissipative deterministic dynamical systems, whose long-term behaviour is dominated by the system attractors. While different types of attractors are observed in different systems, they can generally be classified as either ordered or disordered (chaotic): for example, in time-continuous systems, the former includes fixed points and limit cycles, while the latter is associated with chaotic dynamics. In many systems, there are regions of parameter space corresponding to ordered behaviours, while parameter values in other regions correspond to disordered behaviours.

The separatrices between ordered and disordered regions of parameter space are referred to as *dynamically critical regions*, and the states belonging to these regions are called *dynamically critical states* [10, 11]. Critical regions have also been said to be “at the edge of chaos” [1, 5].

According to the criticality hypothesis, systems that are in critical states should be able to perform better than the others, when interacting with an external *environment*, because ordered systems can be too rigid and incapable to react to changes, while chaotic systems can be too susceptible to small changes to robustly lead to a reliable behaviour [1, 2]. The above statement requires, of course, that there should be a method to evaluate the quality of a system with respect to an environment, i.e., its fitness (biological or artificial).

Studies of artificial evolutionary systems have shown that the criticality hypothesis cannot be true for every kind of system and every kind of task, while it can hold for large classes of systems and tasks [3]. A very important question is whether the hypothesis is valid for biological systems: in this case, the hypothesis translates into the claim that biological evolution has modified the system parameters in such a way that they are likely to be found in critical regions (evolution “towards the edge of chaos”, see Torres-Sosa et al. [12]

and Hidalgo et al. [13]). In biology, two possible amendments to the criticality hypothesis have been proposed:

*Hypothesis 1.* Living systems are likely to be found in the ordered region, but close to the critical boundary [1, 2] in order to be able to react to environmental stimuli and/or, on a longer timescale, to favour evolvability [14].

*Hypothesis 2.* In biological systems, the discontinuities are considerably smoother than the sharp ones observed in physical systems [15] and thus different regimes should still be observed, but they might be separated by a broader transition region.

For recent reviews of the criticality hypothesis, see Mora and Bialek [16] and Roli et al. [17].

Testing the criticality hypothesis in real biological systems, by analyzing the activation behaviour of various genes in cells that are subject to a permanent perturbation, is the topic of the present paper. In order to relate the experimental data with “abstract” dynamical regimes, it is necessary to make use of models: interestingly, our analysis exploits a model that has played a prominent role in the development of complexity science, namely, Random Boolean Networks (RBNs for short, see Kauffman [1, 2, 18], Bastolla and Parisi [19], and Aldana et al. [20]).

In previous papers [21–23], we analyzed the behaviour of the RBN dynamical model for gene regulatory networks, comparing its outcomes with experimental results concerning the change in gene expression levels due to the knockout of single genes in the yeast *S. cerevisiae* [24]. The major outcomes of these studies were (i) the demonstration that even strongly simplified models like RBNs can accurately describe the actual size distribution of the perturbations and (ii) the identification of a way to experimentally test the hypothesis that living beings are preferentially found in dynamically critical states.

The criticality hypothesis can be tested by studying the changes induced by a single-gene knockout (the permanent *silencing* of a single gene) on the expression levels of all the other genes. The genes whose expression levels are significantly modified by the knockout will be referred to as *affected*, and the set of such affected genes will be called an *avalanche* (see also Section 4). The total number of genes that belong to an avalanche is its *size*, and, in those previous works, it has been shown that the size distribution of avalanches provides information about the dynamical regime of the gene regulatory network.

We recently achieved new results concerning both the behaviour of the model and the outcomes of its comparison with experimental data. This has been possible because both new experimental data became available [25], and new theoretical results were derived. In this paper, we describe these new theoretical results and we apply them to the now available larger set of experimental data.

Specifically, we extend the analysis of the RBN model behaviour by providing an analytical formula for the theoretical distribution of noninterfering avalanches (defined in Section 4) of any size, while, in our previous papers, the

analytical distribution was limited to the smallest avalanches, leaving the study of larger ones to simulation only. Moreover, we report numerical results for self-interfering avalanches (also defined in Section 4).

On the empirical side, using the recent experimental data of [25], we obtain a maximum likelihood estimate (MLE) of the Derrida exponent  $\lambda$ ; this is a parameter used to characterize the dynamical regimes of RBNs [26, 27] that fully determines the theoretical avalanche distribution, which makes inference possible from avalanche data. We also assess the uncertainty accompanying our estimate by means of a jackknife estimate of bias and standard error.

In order to compare a Boolean model with real-valued data like those obtained from microarray experiments, it is necessary to introduce a threshold: the smallest change in the expression level of a gene in the model is a switch from 0 to 1, or vice versa, so we need to dichotomize the experimental data. This can be done by considering a gene as modified only if the ratio of its expression level in the knocked-out network to that in the wild-type network is larger than a certain threshold, or smaller than its reciprocal.

In our previous studies, fairly high threshold values ( $\geq 4$ ) had been considered, based upon some heuristic or rule-of-thumb inspired by experimental biologists. Here, taking advantage of the analytical formula for avalanches of any size, we can study how the MLE of  $\lambda$  varies as a function of the threshold. If the plot of MLE vs. threshold had shown a large plateau, one might have claimed that the actual value of the Derrida parameter is scarcely affected by the choice of the threshold in a wide range. Such a plateau does not show up, because the MLE monotonously decreases as the threshold increases, but the picture that emerges from our data analysis is perhaps even more interesting.

With no claim of generalization beyond the data of Kemmeren et al. [25], we have two heuristic arguments that suggest an optimal threshold in a neighbourhood of 3. The first argument is based on the shape of the avalanche size distribution: on general grounds, it is to be expected that smaller avalanches are more frequent than larger ones but, if we use too small a threshold, we observe an initial increase of the empirical frequency as the size grows. There are no plausible reasons for this, but it can be argued that, when the threshold is very small, one actually observes the properties of the noise present in the data; so it might be appropriate to choose a threshold slightly larger than the smallest value that does not display this unphysical behaviour. In this way, we find 2.5 as the minimal threshold value. Note that this argument does not make use of the RBN model, but it is based only upon the data plus a reasonable assumption. The second argument is based on measuring the RBN model fit to the data: here, one observes that the discrepancy between the estimated theoretical distribution and the empirical distribution in the data achieves its minimum at either 3 or 3.5, depending on whether it is measured by the total variation distance or the Kullback–Leibler divergence (see Section 6 for details).

The two arguments point to an interval of plausible threshold values between 2.5 and 3.5, and, while they may not be decisive, they stand together in favour of threshold values within that particular region. With the value 3 chosen

as threshold, the MLE of the Derrida parameter lies in the ordered region, not far from the order-disorder boundary, but with the critical value  $\lambda = 1$  well beyond the margin of error. Furthermore, a supplemental Bayesian analysis assigns negligible posterior probability to the criticality hypothesis. This was one of the scenarios initially suggested by Kauffman [1], and it was also supported by the old analyses from our group [22] and other [28] group. The new data appear to confirm the validity of such a scenario, although clearly no definitive conclusion can be drawn from the analysis of a single organism.

Note that our analysis is not based on a direct analysis of the behaviour in time of the *S. cerevisiae*, which would require dynamical experimental data, but relies on a dynamical model to draw inferences, from static data, on the dynamical regime of the microorganism. As discussed above, it is indeed possible to determine the avalanche size distribution in the model as a function of the Derrida parameter and to compare it with the size distribution that is actually observed in *S. cerevisiae*. This comparison allows one to draw meaningful, albeit indirect, conclusions on the dynamical regime of the biological system.

The outline of the paper is as follows. In Section 2, we provide a general introduction to our modeling framework, while, in Section 3, we specifically review the RBN model. In Section 4, we introduce the definition of noninterfering avalanches and we show that their distribution is correctly approximated by an analytical formula whose derivation is postponed to Section 5. In Section 6, we compare the analytical formula derived in Section 5 with the experimental data collected by Kemmeren et al. [25]. Finally, in Section 7, we summarize our main findings and highlight some open questions.

## 2. Generic Properties of Dynamical Systems

We are interested in stylized models that allow us to explore the generic properties of biological systems, in particular gene regulatory networks. Of course, it is perfectly legitimate to concentrate on a particular hypothesis or a specific genetic-metabolic circuit and to develop a comprehensive model, well-suited to study the details of the system under examination, according to the *system biology* approach [29]. However, these comprehensive models are not suitable for highlighting the general principles of biological organization, which apply to all living beings or at least to their broad classes. We know that there are some principles of this kind, including biological evolution and cellular organization (see chap. 2 of Serra and Villani [30] and the references therein). The study of this kind of phenomena is indeed the challenge of *complex systems* biology [31], which looks for general principles in biological systems.

The criticality hypothesis is a noteworthy example of candidate general principle. As discussed in Section 1, it can be tested by comparing models of biological systems with data, e.g., models of gene regulatory networks with actual gene expression data. The use of data for this purpose is different from the more common use of the same data to infer information about the interactions between specific genes.

In testing the criticality hypothesis, interest lies in global properties of gene expression data, like the avalanche size distribution or some information-theoretic measures [32, 33]. These features of the data are compared to the corresponding predictions from a stylized model, which can be ordered, chaotic, or critical.

A suitable stylized model for gene regulatory networks is an ensemble of networks. In this paper, we obtain an ensemble of networks by generating them from a random system (with random connections and random Boolean functions) while keeping some of its features fixed (e.g., the average number of connections per node). By comparing experimental data to the properties of an ensemble of networks, it is possible to draw inferences on the values of the parameters that define the ensemble (random system). It is therefore possible to compare the distribution of avalanches in real organisms to that in RBN ensembles with different values of the Derrida parameter, and this comparison provides us with information on whether real cells are critical or not. This is a very interesting example of the way in which stylized models can be used to find generic properties, which cannot be read directly from the data, but can be inferred from a comparison between patterns in the data and model predictions.

In recent years, our knowledge about the gene regulatory network of *S. cerevisiae* has considerably improved [25, 34–36] thanks also to initiatives such as DREAM [37, 38]. Nevertheless, we do not know enough details to derive definitive conclusions about its general topology. Indeed, one of the broadest and more accurate reconstructions available of the gene regulatory network of *S. cerevisiae*, summarized by Ma et al. [39] and more comprehensively described in further references therein, shows the presence of very few well-connected genes immersed in a huge “sea” of poorly connected nodes. However, leaving aside the very few well-connected genes, the distribution of connections does not show a very heavy tail, and the subnetwork connecting the well-connected genes is very sparse. Therefore, neither Erdős-Rényi nor scale-free topologies [40, 41] appear to be entirely appropriate. In this paper, we stick to the simplicity of Erdős-Rényi networks, but discuss scale-free alternatives in Section 7.

## 3. Random Boolean Networks

We provide below a synthetic description of the RBN model and its main properties, referring the reader to Kauffman [1, 2], Bastolla and Parisi [19], and Aldana et al. [20], for more detailed accounts. Several variants of the model have been presented and discussed [42] and even used in knockout experiments [43, 44], but we will here restrict our attention to the *classical* model. A classical RBN is a stochastic dynamical system composed of  $r$  nodes, corresponding to genes, which can take either the value 0 (inactive) or the value 1 (active). Let  $X_h(t) \in \{0, 1\}$  be the activation value of node  $h$  at time  $t$ , and let  $X_{1:r}(t) = (X_1(t), X_2(t), \dots, X_r(t))$  be the vector of activation values of all the genes. The relationships between genes are represented by directed links and Boolean functions, which model the response of each node to the values

of its input nodes. In a classical RBN, each node has the same number of incoming connections, and its  $k_{\text{in}}$ , say, input nodes are chosen at random with uniform probability among the remaining  $r - 1$  nodes. In this way, the number of outgoing connections of any given node approximately follows a Poisson distribution with mean  $k_{\text{in}}$  for large  $r$ . The Boolean functions can be chosen in different ways: in this paper, we will only examine some cases where they are chosen at random with uniform probability in a predefined set of allowed transition functions.

In the so-called *quenched* model, both the topology and the Boolean function associated to each node do not change in time. The network dynamics are discrete and synchronous, so fixed points and cycles are the only possible asymptotic states for finite networks. Note that a single RBN can have, and usually has, more than one attractor. The model shows two main dynamical regimes, ordered and disordered, depending upon the degree of connectivity and upon the Boolean functions: typically, the average cycle length grows as a power law of the number of nodes  $r$  in the ordered region and diverges exponentially in the disordered region [1]. The dynamically disordered region also shows sensitive dependence upon the initial conditions, not observed in the ordered one.

It should be mentioned that some interesting analytical results have been obtained by the so-called *annealed* approach [27] in which the topology and the Boolean functions associated to the nodes change at each time step. Several results for annealed networks appear to hold also for the corresponding ensembles of quenched networks, as it has been verified using numerical simulations. Although the annealed approximation may be useful for analytical investigations [20, 45, 46], in this work, we will always be concerned with quenched RBNs, which are obviously closer to real gene regulatory networks, where the topology and the regulatory effects do not undergo continuous changes.

A very important aspect concerns how to determine and measure the dynamical regime of RBNs: while several procedures have been proposed, an interesting and well-known method directly measures the spreading of perturbations through the network. This measure involves two parallel runs of the same system, whose initial states differ for only a small fraction of the nodes. This difference is usually measured by means of the Hamming distance  $H(t)$ , defined as the number of genes that have different activation values on the two runs at time  $t$ ; the measure is performed on many different initial condition realizations, so one actually considers the average value  $\langle H(t) \rangle$ , but we will omit below the somewhat pedantic brackets. If the two runs converge to the same state, i.e.,  $H(t) \rightarrow 0$ , as  $t \rightarrow \infty$ , then the dynamics of the system are robust with respect to small perturbations (a signature of the ordered regime), while if  $H(t)$  grows in time (at least initially), then the system is in the disordered regime. The critical states are those where  $H(t)$  initially remains constant.

If a single node is perturbed, the average number of different nodes at the following time step will be equal to the average number of connections per node times the probability that a node changes value if one of its input changes. This quantity, known as the Derrida parameter

$\lambda$  [26, 27, 47], determines the dynamical regime of the network: if  $\lambda > 1$ , at each step, more nodes will be changed, and the perturbation will grow (chaotic regime); if  $\lambda < 1$ , after a few steps, the perturbation will die out (ordered regime). The value  $\lambda = 1$  is critical (separating chaos from order). The presence of a phase transition in RBNs at a critical value of the Derrida parameter is also confirmed by information-theory-based analyses of their dynamics [17, 33, 48].

In the classical model of RBNs, Boolean functions are often chosen at random among all those with  $k_{\text{in}}$  inputs, but a detailed study of tens of actual genetic control circuits [49] has shown that, in real biological systems, only so-called canalizing functions are found: a function is said to be canalizing if there is at least one value of one of its inputs that uniquely determines its output. It may therefore be interesting to consider models where only canalizing functions are allowed. Moreover, if we associate the value 0 to inactivity, a node that is always 0 will never show its presence, so it may also be interesting to consider models where the null function is excluded [23].

#### 4. Perturbations in Random Boolean Networks

As discussed in Section 1, we can compare what happens in the wild-type and in the knocked-out RBN. In the beginning, a single node (the knocked-out one, also called the root of the perturbation) differs in the two cases, so the initial size of the avalanche is 1. Let  $\mathcal{R}$  be the set of nodes that directly receive input from the root: if no node in  $\mathcal{R}$  changes its value, the avalanche stops at the root, and it turns out to be of size 1; otherwise, the perturbation spreads, and it can later reach nodes that receive inputs from nodes in  $\mathcal{R}$  (possibly through other intermediate nodes). After transients have died out, one can compare the two cases and see how many nodes take different values at least once: these are the nodes that have been *affected* by the perturbation, and their number is the *size* of the *avalanche* induced in that particular network by that particular knockout.

The above definition of avalanche is applicable to simulated data from individual realizations of RBNs. Experimental data obtained by microarray technologies, such as those collected by Hughes et al. [24] and Kemmeren et al. [25] are aggregated data recording gene expression changes in groups of similar cells, and an experimental avalanche will consist of those genes whose expression has changed after a single gene deletion. We will assume that a gene whose expression has changed in the real data corresponds to an affected node in the dynamical model, so we will directly compare the size of the avalanche in the model to the size of the experimental avalanche. For a proper comparison, as already noted in Section 1, it is necessary to dichotomize the experimental data; we will discuss in detail the choice of the dichotomization threshold in Section 6. The comparisons discussed in this section concern only the behaviours of two different models, the wild-type and the knocked-out RBN.

As discussed in detail by Serra et al. [22], some simplifying assumptions can be introduced, which are very well satisfied by the *S. cerevisiae* case.



*Assumption 1.* The network is sparse, i.e., both the number of incoming and outgoing links per node is much smaller than the total number of nodes.

*Assumption 2.* The avalanche size is small with respect to the total number of nodes.

A straightforward consequence of these hypotheses is that the probability that a gene, whose value has changed, feeds back its change into one of its input genes is negligible, and so is the probability that more input nodes of a gene change their values. Under these assumptions, the spreading of an avalanche can be described by a tree, since in these cases, every affected node can be regarded as the root of a subavalanche, which evolves independently of the other subavalanches. We will refer to avalanches of this type as *noninterfering* avalanches, while a *self-interfering* avalanche will be an avalanche where at least one node has two changed inputs (in the knocked-out network with respect to the wild-type one).

In synthesis, if Assumptions 1 and 2 above hold, the number of self-interfering avalanches is very small, and it suffices to consider noninterfering avalanches. From now on, in this paper, the term “avalanche” will be used for noninterfering avalanches unless otherwise explicitly stated. It can be shown [22] that, in the case where self-interfering avalanches are excluded, the probability that the knockout of a randomly chosen node gives rise to an avalanche of size  $v$ , say  $v = 1, 2, \dots$ , depends only upon the probability distribution of the number of outgoing connections. In the classical RBN model considered in this paper, where every node has  $k_{\text{in}}$  incoming links chosen at random with uniform probability among all other  $r - 1$  nodes, when the network is large, as anticipated in Section 3, the outdegree distribution is known to be approximately Poissonian:

$$p_{\text{out}}(k | \alpha) = \frac{\alpha^k}{k!} e^{-\alpha}, \quad k = 0, 1, 2, \dots, \quad (1)$$

where  $p_{\text{out}}(k)$  denotes the probability that a node chosen at random has  $k$  outgoing connections and  $\alpha$  is the average number of outgoing links; note that  $\alpha = k_{\text{in}}$ , because every outgoing link of a node is an input connection for another node. For instance, one finds  $\alpha = 2$  in the simple case of RBNs with two inputs.

More specifically, it can be shown [22] that the avalanche size distribution only depends on the *probability generating function* of the outdegree distribution computed at the probability  $q$  that a node chosen at random is not affected when exactly one of its inputs is affected:

$$g_{\text{out}}(q | \alpha) = \sum_{k=0}^{\infty} q^k \frac{\alpha^k}{k!} e^{-\alpha} = e^{-\alpha(1-q)} \quad (2)$$

in the Poissonian approximation. Serra et al. [22], for RBNs with two inputs, found  $q = 1/2$  when all Boolean functions are allowed,  $q = 4/7$  when only canalizing functions are allowed, or  $q = 7/13$  when the null function is also excluded.

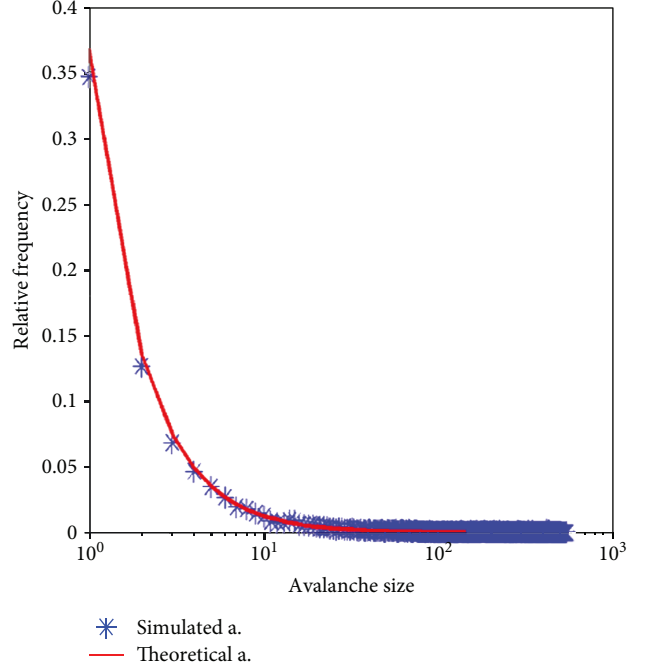


FIGURE 1: Theoretical avalanche distribution (red line), as given by (3), together with the distribution observed in simulations (blue stars) from RBNs with 1000 nodes, 2 inputs per node, and all 16 Boolean function allowed.

It is immediate to check that  $g_{\text{out}}(q | \alpha)$  depends upon  $\alpha$  and  $q$  only through  $\lambda = \alpha(1 - q)$ . This quantity is the product of the average number of connections per node times the probability that a node changes value if one of its input changes, that is, the Derrida parameter introduced in Section 3 (see also Serra et al. [22]). It follows that  $\lambda$  determines the avalanche size distribution, and we will show in Section 5 that

$$p_{\text{ava}}(v | \lambda) = \frac{(\lambda v)^{v-1}}{v!} e^{-\lambda v}, \quad v = 1, 2, 3, \dots, \quad (3)$$

upon knockout of a single gene (for noninterfering avalanches). For instance, in the case of RBNs with two inputs, we have  $\lambda = 1$  (critical value) when all Boolean functions are allowed,  $\lambda = 6/7$  when only canalizing functions are allowed, and  $\lambda = 12/13$  when the null function is also excluded (two instances of ordered regime).

Figure 1 shows that (3) does well approximate the size distribution of avalanches simulated from RBNs with 1000 nodes, two inputs per node, and all Boolean functions allowed (with uniform probability). On the other hand, one can see from Figure 2(b) that the distribution of avalanches in simulated networks with 20 nodes (two inputs per node and only canalizing functions allowed) is not satisfactorily described by (3). This is unsurprising, because we know that Assumptions 1 and 2 above break down for small networks, where self-interference comes into play. Indeed, as illustrated in Figure 2(a), once the simulated avalanches that actually interfere with themselves have been identified and removed, the remaining ones are no longer at odds with the theoretical

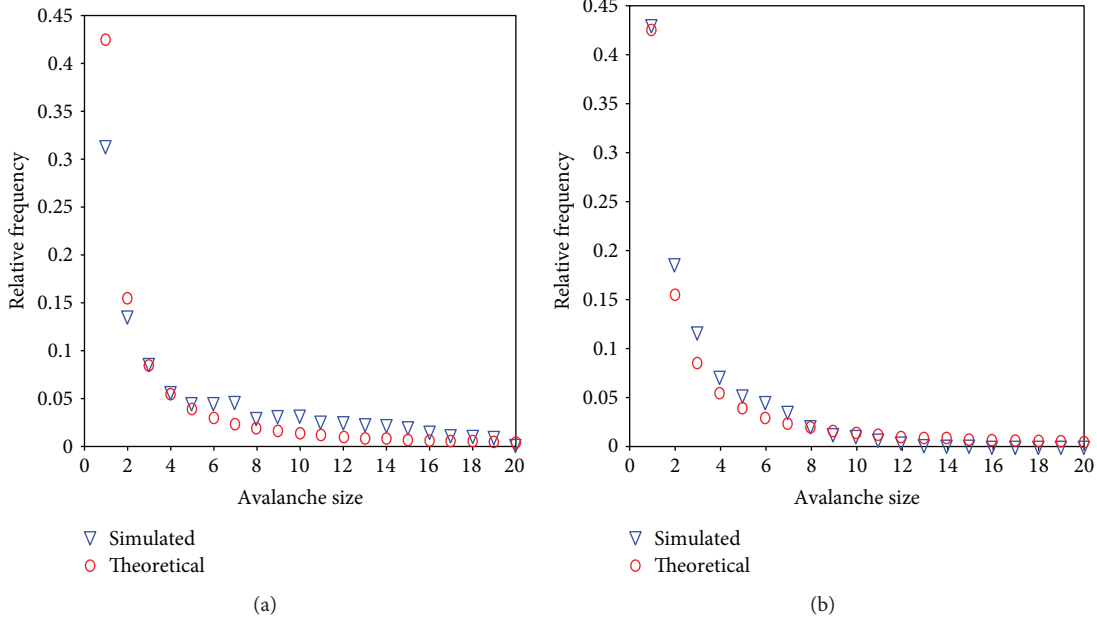


FIGURE 2: Distribution of all avalanches (a) and noninterfering avalanches (b) observed in simulations from RBNs with 20 nodes, 2 inputs per node, and only canalizing functions allowed.

formula (see also Di Stefano et al. [21]). We obtained similar results for networks of different sizes: as it should be expected, *ceteris paribus*, the fraction of self-interfering avalanches is a monotonous decreasing function of the network size  $r$ .

Note that, although the size distribution of noninterfering avalanches given by (3) is not of the power law type, it can closely approximate a power law under some circumstances. Specifically, if we let the network be critical ( $\lambda = 1$ ), then

$$p(v) = \frac{v^{v-1}}{v!} e^{-v}, \quad v = 1, 2, 3, \dots, \quad (4)$$

where we have simplified our notation by writing  $p(v)$  in place of  $p_{\text{ava}}(v | 1)$ . Then, if the avalanche is not too small, one can insert Stirling's approximation  $v! \approx \sqrt{2\pi v} (v/e)^v$  in (4) and thus obtain

$$p(v) \approx \frac{1}{\sqrt{2\pi}} v^{-3/2}, \quad v = 1, 2, 3, \dots, \quad (5)$$

i.e., a power law with slope 1.5. Interestingly, the approximation is quite good: the error due to Stirling's formula is 8% for the smallest avalanches ( $v=1$ ), and it rapidly falls to 4% for avalanches of size two, to 2% for avalanches of size four, to 1% for avalanches of size six, and below 1% for avalanches whose size is nine or larger.

It is also important to note that the size distribution of avalanches is very similar to the outdegree distribution of the Genetic Perturbation Network (GPN) in the paper by Kemmeren et al. [25], whose experimental data about changes in gene expression levels will be compared with our theoretical predictions in Section 6. Indeed, the GPN of Kemmeren et al. [25] includes a link from each knocked-

out gene to all those genes whose expression levels have significantly changed. Hence, in the GPN, a knocked-out gene is directly linked to all genes affected by its avalanche.

We stress that the GPN is not the gene regulatory network, which we model as RBN, because the latter includes only direct causal links. Suppose, for example, that gene  $a$  is knocked out and that it has output connections to genes  $b$  and  $c$ . Suppose also that gene  $b$  has output connections to genes  $d$  and  $e$ . Finally, for the sake of definiteness, suppose that  $b, c, d$ , and  $e$  are all affected when  $a$  is knocked out. Then, in the GPN node,  $a$  will have four links (to  $b, c, d$ , and  $e$ ) while in the RBN, it will only have two links (to  $b$  and  $c$ ) although there will be two further links in the network (from  $b$  to  $d$  and  $e$ ). It is apparent that the outdegree distribution of the RBN is different from the outdegree distribution of the GPN.

Kemmeren et al. [25] claim that the GPN of *S. cerevisiae* exhibits a power law outdegree distribution. In light of the above described approximation, their claim is compatible with adopting a critical RBN as gene regulatory network. Indeed, using the same thresholds for differential expression as Kemmeren et al. [25], we would obtain in Section 6 an estimate of the Derrida parameter very close to one ( $\hat{\lambda} = 0.98$ ). In fact, our study of threshold choice will lead us to a different result.

## 5. Theoretical Avalanche Size Distribution

We now derive the probability mass function of the avalanche size in a large RBN. Specifically, the RBN should be large enough for the number of outgoing connections of any given node to follow a Poisson distribution and for the avalanche to progress without interfering with itself. We remark that both conditions can be satisfied only in an approximate sense, which is the reason why we regard the

resulting distribution as theoretical. The same distribution was derived by R  m   et al. [28], assuming the annealed approximation, but following a less elementary line of proof.

Under the noninterference condition, each node in the avalanche has a unique parent that determines its first deviation from the unperturbed trajectory, except for the root node (the first perturbed node). The avalanche size  $V$  therefore satisfies the equation  $V = 1 + W_1 + \dots + W_v$ , where  $W_v$  is the number of affected nodes that have as parent the  $v$ -th perturbed node; note that the actual ordering of nodes perturbed at the same time is immaterial. Clearly, we have  $1 + W_1 + \dots + W_v > v$  for all  $v < V$ , and  $W_V = 0$ , so that  $V = \inf \{v \geq 1 : 1 + W_1 + \dots + W_v = v\}$ . If we define  $Z_v = W_v - 1$ , for all  $v \geq 1$ , we obtain  $V = \inf \{v \geq 1 : 1 + Z_1 + \dots + Z_v = 0\}$ .

Since  $W_1, W_2, \dots$ , hence  $Z_1, Z_2, \dots$ , are independent and identically distributed in the RBN model, we have shown that the avalanche size equals the hitting time of the origin for a random walk on the integers starting at 1 and having steps bounded from below by  $-1$ . For such a random walk, the following result (known as the *hitting time theorem*) holds with an elementary proof [50]:

$$\Pr(V = v) = \frac{1}{v} \Pr(1 + Z_1 + \dots + Z_v = 0), \quad v = 1, 2, 3, \dots, \quad (6)$$

where the left-hand side of (6) defines the probability mass function  $p_{\text{ava}}(v | \lambda)$  of  $V$ , and we will turn the right-hand side of (6) into the closed-form expression appearing in the right-hand side of (3).

Let  $Y_v$  be the number of outgoing connections of node  $v$ . Recall that  $Y_v$  follows a Poisson distribution with probability mass function (1). Since exactly  $j$  out of  $k$  outputs must be affected to obtain  $W_v = j$  when  $Y_v = k$ ,  $k = 0, 1, 2, \dots$ , we can write

$$\Pr(W_v = j | Y_v = k) = \frac{k!}{j!(k-j)!} (1-q)^j q^{k-j}, \quad j = 0, 1, \dots, k. \quad (7)$$

Then, the law of total probability gives us

$$\begin{aligned} \Pr(W_v = j) &= \sum_{k=j}^{\infty} \frac{k!}{j!(k-j)!} (1-q)^j q^{k-j} p_{\text{out}}(k | \alpha) \\ &= \frac{(1-q)^j}{j!} \sum_{k=j}^{\infty} \frac{q^{k-j}}{(k-j)!} \alpha^k e^{-\alpha} \\ &= \frac{\alpha^j (1-q)^j}{j!} e^{-\alpha+q\alpha} \end{aligned} \quad (8)$$

for all  $j = 0, 1, 2, \dots$ , so that  $W_v$  follows a Poisson distribution with mean equal to the Derrida parameter  $\lambda = \alpha(1-q)$ .

Finally, it is a standard result that  $W_1 + \dots + W_v$  follows the Poisson distribution with mean  $v\lambda$ . This allows us to compute the right-hand side of (6) as

$$\frac{1}{v} \Pr(W_1 + \dots + W_v = v-1) = \frac{(\lambda v)^{v-1}}{v!} e^{-\lambda v}, \quad v = 1, 2, 3, \dots, \quad (9)$$

which is the desired closed-form expression for  $p_{\text{ava}}(v | \lambda)$  in (3).

We remark that (3) is a valid probability mass function for all  $\lambda > 0$ , if complemented by  $p_{\text{ava}}(\infty | \lambda) = 1 - \sum_{v=1}^{\infty} p_{\text{ava}}(v | \lambda) = \Pr(V = \infty)$ ; this quantity is strictly positive, hence nonignorable, if and only if  $\lambda > 1$  (chaotic regime). The last assertion follows from the fact that  $V = 1 + W_1 + \dots + W_V$  can be seen as the total progeny of a single ancestor branching process having Poisson distributed offsprings with mean  $\lambda$ . The same representation shows that the mean of  $V$  is given by  $\sum_{n=0}^{\infty} \lambda^n$  (sum of the means of all generations in the branching process) which is finite and equals  $1/(1-\lambda)$  if  $\lambda < 1$  (see, for example, Harris [51], chap. I).

## 6. Comparison with Experimental Data

We analyze genome-wide mRNA expression data from a battery of gene deletion experiments on the yeast *S. cerevisiae*. Specifically, we analyze the data collected by Kemmeren et al. [25]. These authors measured the expression change of  $r = 6182$  genes in  $n = 1484$  mutant yeast strains carrying a single gene deletion. For each gene and mutant, they produced two numbers: a fold change and a  $P$  value. Let  $m_{hi}$  be the binary logarithm of the fold change for gene  $h$  and mutant  $i$ , and denote by  $p_{hi}$  the corresponding  $P$  value. We refer the reader to Kemmeren et al. [25] for all information on the process leading to the quantities  $m_{hi}$  and  $p_{hi}$ ,  $h = 1, \dots, r$ ,  $i = 1, \dots, n$ , which we regard here as our raw data. We use the software R [52] for all data analyses, both numerical and graphical.

In order to adopt the RBN model as the data generating mechanism, we need to dichotomize the data. This can be done by declaring a gene differentially expressed in a given mutant (with respect to the wild-type strain) when the corresponding fold change is large (in absolute value) and the corresponding  $P$  value is small. Let  $e_{hi} = 1$  if  $|m_{hi}| \geq \log_2(T)$  and  $p_{hi} \leq P$ , or  $e_{hi} = 0$  otherwise, so that  $e_{hi}$  indicates differential expression of gene  $h$  in mutant  $i$ . The size of the avalanche originated by gene deletion in mutant  $i$  is thus given by  $v_i = \sum_{h=1}^r e_{hi}$ , for all  $i = 1, \dots, n$ , and the empirical distribution of  $v_1, \dots, v_n$  can be compared with the avalanche size distribution in the RBN (for any given value of the Derrida parameter  $\lambda$ ).

A comparison along the above described lines was carried out by Serra et al. [22, 23] on a smaller dataset collected by Hughes et al. [24]. In a pioneering study, the latter authors had measured expression changes genome-wide for 276 deletion mutants of *S. cerevisiae*. Serra et al. [23] restricted their attention to fourfold or larger changes ( $T = 4$ ) and ignored the  $P$  values ( $P = 1$ ) to focus on incontrovertible changes without relying on any null hypothesis. They found a good graphical agreement between the size distribution of experimental avalanches and the size distribution of avalanches

simulated from RBNs with two inputs allowing only canalizing functions ( $\lambda = 6/7$ ). Serra et al. [22] let  $T$  vary between 2 and 15 and found that the choice  $T = 7$  was optimal in terms of  $\chi^2$  distance between the two distributions (for all three types of RBNs with two inputs). With this choice, using the same optimality criterion, but replacing the simulated distribution with the theoretical one, whose analytical expression they had derived up to size six, and dropping observed avalanches of size seven or more, they estimated  $\lambda = 0.84$  (quite close to  $6/7 \approx 0.86$ ). Note that all strictly positive values of  $\lambda$  are of interest, in principle, because RBNs are a very abstract representation of gene interaction networks, and the case with two inputs is just the simplest possible scenario (not quite a realistic one).

We here compare the full theoretical distribution derived in Section 5 to the empirical avalanche size distribution in the larger dataset collected by Kemmeren et al. [25]. Specifically, we analyze the avalanche sizes  $v_1, \dots, v_n$  obtained with a given choice of  $T$  and  $P$  using the statistical model with likelihood function:

$$L(\lambda) = \prod_{i=1}^n p_{\text{ava}}(v_i | \lambda) = \left\{ \prod_{i=1}^n \frac{v_i^{v_i-1}}{v_i!} \right\} \lambda^{n(\bar{v}-1)} e^{-n\bar{v}\lambda}, \quad \lambda > 0, \quad (10)$$

where  $\bar{v} = n^{-1} \sum_{i=1}^n v_i$  is the observed mean avalanche size. Equation (10) gives the theoretical probability of observing  $v_1, \dots, v_n$  as avalanche sizes, in  $n$  independent RBNs, as a function of the Derrida parameter. We study the sensitivity of our results to the choice of  $T$  and  $P$  with methods discussed later in this section.

It is easily seen that (10) admits a unique maximum for  $\hat{\lambda} = 1 - 1/\bar{v}$ , which is therefore the maximum likelihood estimate of  $\lambda$ . Note that  $\hat{\lambda} < 1$  for all possible values of  $v_1, \dots, v_n$ , because we cannot observe infinite avalanche sizes. This means that we can never learn a Derrida parameter larger than one. However, if  $\lambda > 1$ , we would sooner or later observe an infinite avalanche size and this would rule out all  $\lambda \leq 1$ , while for  $\lambda = 1$ , we will observe  $\bar{v}$  tending to infinity as  $n$  grows and this will give us a value of  $\hat{\lambda}$  close to 1 for large  $n$  (albeit always strictly smaller than 1). In practice, we need to check that the observed avalanches are small with respect to the network and to assess the bias associated to the estimate  $\hat{\lambda}$ . We carry out the first task by comparing  $v_{(n)} = \max\{v_1, \dots, v_n\}$  to the number of genes  $r$ , and the second task by means of the jackknife method [53, 54]. We also use this method to assess the precision of  $\hat{\lambda}$  by computing its jackknife standard error.

Table 1 reports our results with  $P = 1$ , that is, ignoring  $P$  values, and  $T$  varying from twofold to fourfold. Since some of the observed avalanches have size zero, following Serra et al. [22], we merge them with those of size one. Alternatively, they can be dropped: if we cannot even record a change for the deleted gene, it may be wiser to ignore the whole deletion experiment. Table 2 reports our results with this alternative approach: it can be seen that

TABLE 1: Results with  $P = 1$  (zeros merged with ones).

T	$n$	vMax	vMean	MLE	JBias	JStdErr	KLD	TVD
2	1484	741	23.4	0.96	-0.0002	0.0027	0.258	0.276
2.25	1484	625	16.4	0.94	-0.0003	0.0041	0.149	0.192
2.5	1484	530	12.2	0.92	-0.0004	0.0058	0.099	0.128
2.75	1484	485	9.6	0.90	-0.0006	0.0078	0.081	0.106
3	1484	440	7.8	0.87	-0.0007	0.0099	0.083	0.130
3.25	1484	396	6.5	0.85	-0.0009	0.0121	0.092	0.150
3.5	1484	356	5.6	0.82	-0.0012	0.0146	0.109	0.188
3.75	1484	331	4.8	0.79	-0.0014	0.0173	0.132	0.217
4	1484	306	4.3	0.77	-0.0016	0.0199	0.154	0.237

TABLE 2: Results with  $P = 1$  (zeros dropped).

T	$n$	vMax	vMean	MLE	JBias	JStdErr	KLD	TVD
2	1445	741	24.0	0.96	-0.0002	0.0027	0.292	0.298
2.25	1414	625	17.2	0.94	-0.0003	0.0039	0.182	0.224
2.5	1364	530	13.2	0.92	-0.0004	0.0054	0.128	0.167
2.75	1322	485	10.6	0.91	-0.0005	0.0070	0.094	0.118
3	1278	440	8.9	0.89	-0.0007	0.0088	0.081	0.097
3.25	1231	396	7.7	0.87	-0.0008	0.0105	0.075	0.110
3.5	1182	356	6.7	0.85	-0.0010	0.0125	0.072	0.113
3.75	1154	331	5.9	0.83	-0.0012	0.0147	0.084	0.147
4	1110	306	5.4	0.81	-0.0014	0.0167	0.096	0.167

here  $n$  varies with  $T$ , which allows one to appraise the magnitude of the phenomenon. We see from either Table 1 or Table 2 that all observed avalanches are small with respect to the total number of genes in the network:  $v_{(n)}$  is at most 741, while  $r = 6182$ . We also see that, in both tables, the jackknife method estimates the bias as negligible with respect to the standard error (by an order of magnitude). However, it appears that the value of  $\hat{\lambda}$  strongly depends on the choice of  $T$ : it ranges from 0.96 down to 0.77 (Table 1) or 0.81 (Table 2). There is no robust estimate of the Derrida parameter with respect to threshold choice, even though the dependence of  $\hat{\lambda}$  on  $T$  is less pronounced when dropping zeros than when merging them with ones, and we therefore resort to the two heuristic arguments anticipated in Section 1.

Figure 3, focussing on the smallest thresholds, plots the difference between the number of size one avalanches and the number of size two avalanches as a function of the chosen threshold. We read Figure 3 as suggesting that the minimal threshold fully reverting the unphysical inversion between the two empirical frequencies should be a value slightly higher than 2.3. Then, we compare the results obtained with different values of  $T$  by assessing how good is the fit of the statistical model for the different data dichotomizations. To this aim, we gauge how close is the estimated distribution, represented by the theoretical probability mass function



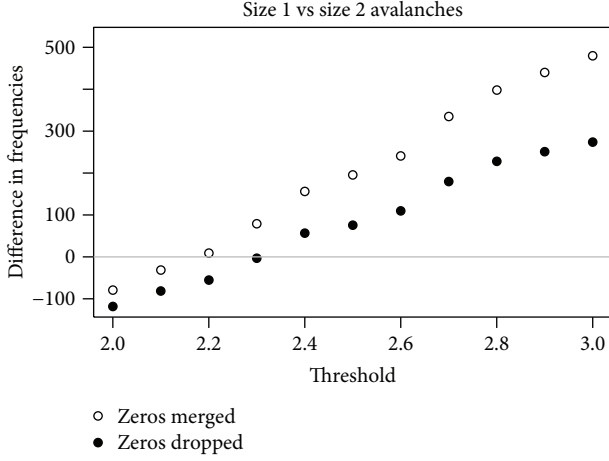


FIGURE 3: Difference in frequencies between avalanches of size 1 and avalanches of size 2 as a function of the dichotomization threshold (ignoring  $P$  values).

with Derrida parameter  $\hat{\lambda}$ , to the empirical distribution of  $v_1, \dots, v_n$ , represented by the probability mass function.

$$p_{\text{emp}}(v) = \frac{1}{n} \sum_{i=1}^n I_{v_i}(v), \quad v = 1, 2, 3, \dots, \quad (11)$$

where  $I_{v_i}(v)$  indicates whether we have observed  $v_i = v$  or not:  $I_{v_i}(v) = 1$  if  $v_i = v$ , or  $I_{v_i}(v) = 0$  otherwise. There are different ways to gauge (lack of) closeness between two probability mass functions, and we select two on grounds of interpretability.

The Kullback–Leibler divergence, or relative entropy ([55], chap. 2), between the empirical mass function  $p_{\text{emp}}$  and the estimated theoretical mass function  $\hat{p}_{\text{ava}} = p_{\text{ava}}(\cdot | \hat{\lambda})$  is given by

$$\text{KLD}(p_{\text{emp}} \| \hat{p}_{\text{ava}}) = \sum_{v=1}^{\infty} p_{\text{emp}}(v) \log \frac{p_{\text{emp}}(v)}{\hat{p}_{\text{ava}}(v)}. \quad (12)$$

It measures, on an information scale, the inefficiency of assuming that the mass function is  $\hat{p}_{\text{ava}}$  when it is  $p_{\text{emp}}$ . We have  $\text{KLD}(p_{\text{emp}} \| p) \geq 0$ , for all probability mass function  $p$  on the natural numbers, and  $\text{KLD}(p_{\text{emp}} \| p) = 0$  if and only if  $p \equiv p_{\text{emp}}$ . It is worth noting that  $\hat{p}_{\text{ava}}$  minimizes  $\text{KLD}(p_{\text{emp}} \| p_{\lambda})$  over all  $p_{\lambda} = p_{\text{ava}}(\cdot, \lambda)$  in the statistical model, because  $\text{KLD}(p_{\text{emp}} \| p_{\lambda})$  coincides with the log-likelihood up to a negative affine transformation. Therefore, we can see  $\text{KLD}(p_{\text{emp}} \| \hat{p}_{\text{ava}})$  as a measure of the inefficiency of using the statistical model in place of the empirical distribution.

The total variation distance, or distance in variation ([56], chap. 4), between  $p_{\text{emp}}$  and  $\hat{p}_{\text{ava}}$  is defined as

$$\text{TVD}(p_{\text{emp}}, \hat{p}_{\text{ava}}) = \frac{1}{2} \sum_{v=1}^{\infty} |p_{\text{emp}}(v) - \hat{p}_{\text{ava}}(v)|. \quad (13)$$

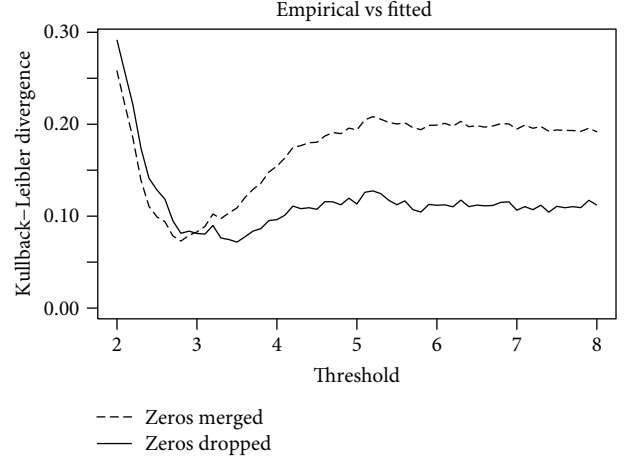


FIGURE 4: Kullback–Leibler divergence between the empirical and estimated avalanche distributions as a function of the dichotomization threshold (ignoring  $P$  values).

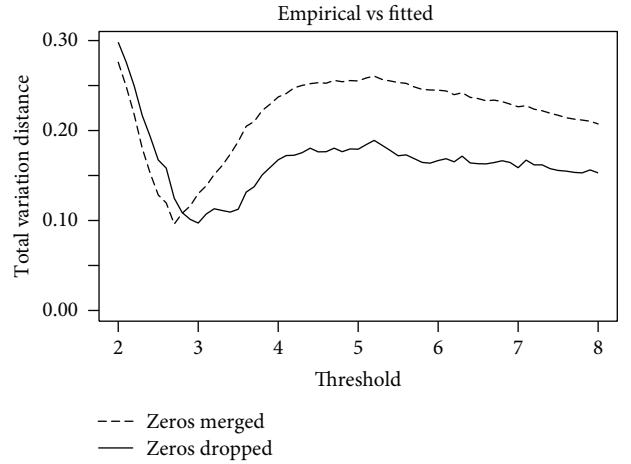


FIGURE 5: Total variation distance between the empirical and estimated avalanche distributions as a function of the dichotomization threshold (ignoring  $P$  values).

It equals the supremum over all subsets of the natural numbers for the absolute difference between the probability computed using  $p_{\text{emp}}$  and that computed using  $\hat{p}_{\text{ava}}$ .  $\text{TVD}(p_{\text{emp}}, \hat{p}_{\text{ava}})$  therefore provides a sharp bound on the worst case error committed by replacing  $p_{\text{emp}}$  with  $\hat{p}_{\text{ava}}$  (or vice versa); since its maximum possible value is 1, we can read it as a percentage, even though it is not a measure of relative error.

Figure 4 plots  $\text{KLD}(p_{\text{emp}} \| \hat{p}_{\text{ava}})$  as a function of the chosen threshold  $T$ . The fit is better when zeros are dropped, than when they are merged with ones, but for the smallest thresholds, where it is either unsatisfactory in both cases or comparable in the two cases. A similar scenario is depicted in Figure 5, where  $\text{TVD}(p_{\text{emp}}, \hat{p}_{\text{ava}})$  is plotted in place of  $\text{KLD}(p_{\text{emp}} \| \hat{p}_{\text{ava}})$ . Therefore, in the following, we stick to dropping zeros, and we refer to the results in Table 2; recall that this choice also results in a more robust MLE. The optimal threshold choice in terms of Kullback–Leibler divergence is

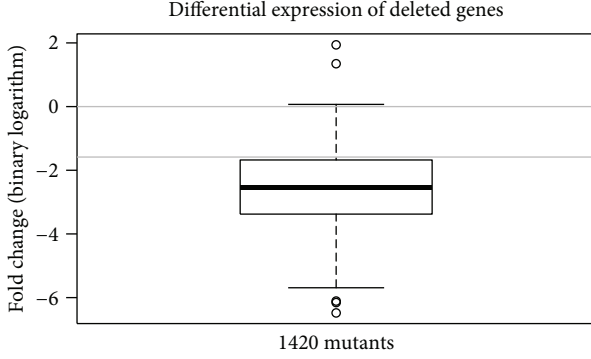


FIGURE 6: Fold change distribution for deleted genes; horizontal gray lines at zero (no change) and  $-\log_2(3) \approx -1.58$  (optimal threshold).

TABLE 3: Results with  $P = 0.01$  (zeros dropped).

T	$n$	vMax	vMean	MLE	JBias	JStdErr	KLD	TVD
2	1442	741	23.9	0.96	-0.0002	0.0027	0.267	0.273
2.25	1411	625	17.1	0.94	-0.0003	0.0040	0.166	0.203
2.5	1361	530	13.2	0.92	-0.0004	0.0054	0.123	0.160
2.75	1320	485	10.6	0.91	-0.0005	0.0070	0.091	0.114
3	1278	440	8.9	0.89	-0.0007	0.0088	0.080	0.097
3.25	1231	396	7.7	0.87	-0.0008	0.0105	0.074	0.110
3.5	1182	356	6.7	0.85	-0.0010	0.0125	0.072	0.114
3.75	1154	331	5.9	0.83	-0.0012	0.0147	0.085	0.148
4	1110	306	5.4	0.81	-0.0014	0.0167	0.097	0.168

given by  $T = 3.5$ , while the optimal threshold in terms of total variation distance is  $T = 3$ . Keeping in mind our first heuristic argument, we select  $T = 3$  as our best choice, and we note that it provides a good fit under both criteria of discrepancy; in particular, with  $T = 3$ , we get a less than 10% worst case error by replacing the empirical distribution with the estimated theoretical one.

The considerable difference between our best choice  $T = 3$  and the optimal choice  $T = 7$  found by Serra et al. [22] can be explained in terms of the remarkable improvements in measurement technology which occurred between the work of Hughes et al. [24] and that of Kemmeren et al. [25]. As a double check on the implications of our choice, we draw in Figure 6 a boxplot of fold changes obtained for deleted genes in the corresponding deletion experiments. It is clear from Figure 6 that the central part of the distribution is below the optimal threshold.

Table 3 has the same structure as Table 2, but it is based on  $P = 0.01$ . This is intended as a high significance threshold, and since the results in Table 3 are very similar to those in Table 2, we conclude that the choice of  $P$  is not critical. This shows that to identify large changes, we need not rely on any null hypothesis, and we therefore stick to our original choice  $P = 1$ . As an aside, we remark that Kemmeren et al. [25] defined differential expression as  $T = 1.7$  and  $P = 0.05$  in their analyses. This choice, which was presented as a

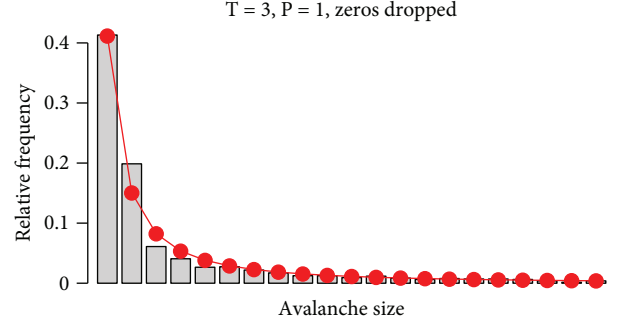


FIGURE 7: Empirical (gray bars) and estimated theoretical (red bullets) distributions with our final data analytical choices.

stringent form of thresholding, aimed at focussing on robust changes more likely to be biologically meaningful, does not lead to a good fit of our model. Of course, in general, different choices are appropriate to different analyses, but maybe our results can offer some insight here.

Letting  $T = 3$  (with  $P = 1$  and dropping zeros), we obtain  $\hat{\lambda} = 0.89 \pm 0.02$  (two standard errors, see Figure 7 for a graphical representation of the corresponding fit). The estimated value of the Derrida parameter is thus intermediate between  $\lambda_1 = 6/7 \approx 0.86$  (theoretical value for RBNs with two inputs allowing only canalizing functions) and  $\lambda_2 = 12/13 \approx 0.92$  (theoretical value for RBNs with two inputs also excluding the null function). We compare the support given by the data to these two special values, and to the third special value  $\lambda_3 = 1$  (theoretical value for RBNs with two inputs allowing all Boolean functions and critical value), by means of an elementary Bayesian model.

We introduce a random variable  $\Lambda$  to represent the unknown value of the Derrida parameter, and we assign  $\Pr(\Lambda = \lambda_1) = \Pr(\Lambda = \lambda_2) = \Pr(\Lambda = \lambda_3) = 1/3$  as prior probabilities (restricting our attention to the special values of interest). Then, we use an elementary version of Bayes formula to compute

$$\Pr(\Lambda = \lambda_j | V_{1:n} = v_{1:n}) = \frac{\Pr(V_{1:n} = v_{1:n} | \Lambda = \lambda_j) \Pr(\Lambda = \lambda_j)}{\sum_{k=1}^3 \Pr(V_{1:n} = v_{1:n} | \Lambda = \lambda_k) \Pr(\Lambda = \lambda_k)}, \quad (14)$$

for  $j = 1, 2, 3$ , where  $V_{1:n} = v_{1:n}$  is a shorthand for  $V_1 = v_1, \dots, V_n = v_n$  (the observed event) and  $\Pr(V_{1:n} = v_{1:n} | \Lambda = \lambda) = L(\lambda)$  is the likelihood of the value  $\lambda$ . In this way, we find the following posterior probabilities: about 85% for  $\lambda_1 = 6/7$ , about 15% for  $\lambda_2 = 12/13$ , and about zero for  $\lambda_3 = 1$ .

The above results suggest RBNs with two inputs allowing only canalizing functions (including the null function) as the data generating mechanism for the avalanches observed by Kemmeren et al. [25]. These are subcritical RBNs, while critical RBNs are not at all supported by the data: the critical value  $\lambda_3 = 1$  receives a negligible posterior probability in our model. However, one might argue that  $\lambda_1 = 6/7$  and  $\lambda_2 = 12/13$  are two lucky subcritical values for the Derrida parameter, selected by the arbitrary choice of RBNs with

two inputs, and that the comparison is not fair, because critical RBNs have no tuning parameter. We can fairly assess whether  $\lambda = 1$  or  $\lambda \neq 1$  by means of a more elaborate Bayesian model. This also results in an extremely small posterior probability of criticality (see Appendix for details).

## 7. Conclusions

We have shown in this paper how it is possible to draw inferences about the dynamical regime of a genetic network from the study of data that do not change in time. As discussed in the previous sections, this can be achieved by resorting to a dynamical model of the network and by showing that the size distribution of avalanches depends upon the same parameter that determines the dynamical regime. Our estimate of the Derrida parameter is therefore indirect, and yet it provides a strong support to the conclusion that the gene network of *S. cerevisiae* is slightly ordered. Similar conclusions had been found in HeLa cells [11] by an entirely different method, based upon the use of Lempel-Ziv complexity.

The experimental data about the effects of gene perturbations (and therefore about the size of avalanches) do not come from a single cell, but rather from a group of nonsynchronized cells, and the measurements were taken much later than the knockouts had been performed. The sizes of the experimentally measured avalanches should therefore be compared with those of the fully developed simulated avalanches, as it has been done in this paper.

The application of the method is limited to experiments based on single gene knockouts, because, in this case, the size of the initial perturbation of the network activation values is known. Other information about the dynamical regime of biological networks might be obtained by analyzing time-course data, that is, by studying the time evolution of avalanche sizes. In this way, it might be possible to use also data from more general perturbations, which might directly affect the expression of several genes, like, e.g., those produced by exposing cell cultures to contaminants or drugs. However, classical RBNs are not well-suited to deal with time courses because of their synchronous update, which amounts to assuming that all the mRNA is completely degraded between two successive time steps; otherwise, the state at time  $t + 1$  would not depend only upon the state at time  $t$  but also upon previous time steps, as discussed by Graudenzi et al. [43]. Since different mRNAs may have different decay constants, this assumption seems questionable. To circumvent this problem, it is possible to introduce memory in the model (see, e.g., Graudenzi et al. [43, 57] and further references therein). This modification also affects the very definition of an avalanche and is a topic for further research [44].

Note that we used fully random Boolean networks, where there is nothing specific about *S. cerevisiae*. The implicit assumption is that the distribution of avalanche sizes be a generic property, so that a generic model can capture it; similar remarks apply to the abovementioned work on HeLa cells [11]. The assumption seems quite justified a posteriori, by the good agreement with experimental data.

Other cases exist where a Boolean model has been used to describe specific gene regulatory networks, where

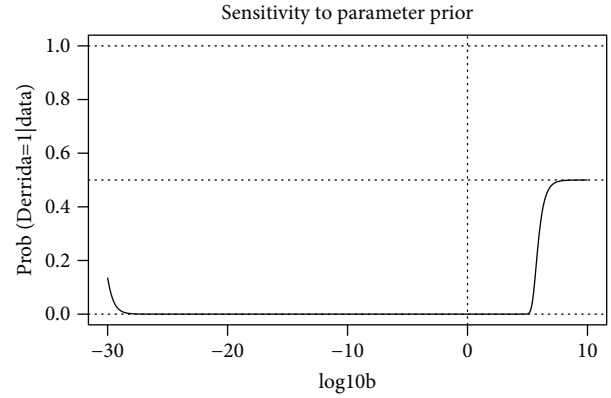


FIGURE 8: Posterior probability of criticality ( $\lambda = 1$ ) as a function of prior conditional precision (inverse of prior conditional variance) given  $\lambda \neq 1$ .

the connections are supposed to be known. There are a priori some caveats to be taken into account, since the regulatory influences that are known might be only a subset of a wider, largely unknown network. Having said this, it should be stressed that, in some cases, these “specific” networks were found to be close to criticality [58].

An important point concerns the actual degree distribution of the nodes of *S. cerevisiae*. We have shown elsewhere [22] that, in the case of noninterfering avalanches, their size distribution does not depend upon the exact shape of the indegree distribution of the gene regulatory network. It can therefore be assumed for simplicity that the indegree is identical for every node. On the other hand, the avalanche size distribution can and does depend upon the shape of the out-degree distribution. As discussed at the end of Section 2, recent experimental work shows that the genetic regulatory network of *S. cerevisiae* is not of the Erdős-Rényi type, nor it can be claimed to be of the scale-free type. Some networks built upon experimental data, like the GPN of Kemmeren et al. [25], appear to be scale-free. However, as discussed at the end of Section 4, they do not represent the causal influences between the genes (see also Wagner [59]). The safest conclusion is that we do not have a reliable knowledge of the topology of the (causal) genetic regulatory network, and we have lessons to learn from simplified models like the RBN discussed in this paper, or scale-free networks, or others. This is also consistent with the general approach of looking for generic properties, rather than specific features.

Scale-free networks have been discussed in depth in a previous paper by our group [60], where it was shown that the avalanche size distribution on such a network can closely approximate the one on an Erdős-Rényi network, and also the available experimental data. However, in order to properly describe the distribution of the smaller (and more frequent) avalanches, it was necessary to choose the parameters in such a way that also very large avalanches were possible. These very large avalanches were not observed in the data available at that time [24], but those data referred to a limited number of knockouts, and one could not exclude that giant avalanches might have been observed if a larger set of knockouts had been performed. We plan to compare in

further work the distribution of avalanches in scale-free Boolean networks with the wider data set of Kemmeren et al. [25], investigating its dependency upon the value of the Derrida parameter.

A further remark is in order. As observed by Kauffman [1, 2], the line of reasoning that leads to the criticality hypothesis applies to whole organisms, in the case of multicellular organisms, rather than to isolated cells. *S. cerevisiae* is a unicellular organism, yet it is known that it can live in colonies [61], a form of organization that can be considered as intermediate between unicellularity and true multicellularity. Indeed, in colonies, different cells can have different behaviours and different gene activation patterns, depending, e.g., upon the position they occupy and the access they have to nutrients. The requirement of criticality for *S. cerevisiae* might therefore be proposed for colonies, rather than for isolated cells, and, in this case, it would be important to relate the dynamical behaviour of a colony to that of a single cell. Some models have been proposed to explore this relationship [62–66].

In conclusion, let us remark that this study is limited to a single organism, so no general claim about criticality can be advanced. However, we have shown a way to analyze the dynamics of biological genetic networks that might be easily generalized and thus be applied to a much wider set of different records.

## Appendix

### Bayesian Test of Criticality

In order to assess whether  $\lambda = 1$  or  $\lambda \neq 1$  without focussing on any other particular value of the Derrida parameter, we specify the prior distribution of  $\Lambda$  as

$$\Pr(s \leq \Lambda \leq t) = \frac{1}{2} I_{[s,t]}(1) + \frac{1}{2} \int_s^t \pi(\lambda) d\lambda, \quad 0 < s < t < \infty, \quad (\text{A.1})$$

where  $I_{[s,t]}(1) = 1$  if  $s \leq 1 \leq t$ , or  $I_{[s,t]}(1) = 0$  otherwise, and  $\pi$  is a probability density function on the positive real numbers. In this way, we have  $\Pr(\Lambda = 1) = 1/2 = \Pr(\Lambda \neq 1)$ , while  $\pi$  is the conditional probability density of  $\Lambda$  given  $\Lambda \neq 1$ . Then, a more elaborate version of Bayes formula gives us

$$\Pr(\Lambda = 1 \mid V_{1:n} = v_{1:n}) = \frac{(1/2)L(1)}{(1/2)L(1) + (1/2)\int_0^\infty L(\lambda)\pi(\lambda)d\lambda}, \quad 0 < a < b < \infty, \quad (\text{A.2})$$

where  $L(\lambda) = \Pr(V_{1:n} = v_{1:n} \mid \Lambda = \lambda)$  is the likelihood (10). On grounds of simplicity, aiming at computing the integral in the denominator of (A.2), we let

$$\pi(\lambda) = \frac{b^a}{\Gamma(a)} \lambda^{a-1} e^{-b\lambda}, \quad \lambda > 0, \quad (\text{A.3})$$

for some  $a, b > 0$ , where  $\Gamma(a) = \int_0^\infty \lambda^{a-1} e^{-\lambda} d\lambda$ ,  $a > 0$ , is the well-known gamma function. With this choice, we obtain

$$\int_0^\infty L(\lambda)\pi(\lambda)d\lambda = \left\{ \prod_{i=1}^n \frac{v_i^{v_i-1}}{v_i!} \right\} \frac{b^a}{\Gamma(a)} \frac{\Gamma(a_*)}{b_*^a}, \quad (\text{A.4})$$

where  $a_* = a + n(\bar{v} - 1)$  and  $b_* = b + n\bar{v}$ . We can therefore use (A.2) to compute the posterior probability of criticality, as soon as we specify  $a, b > 0$ . Since the conditional mean of  $\Lambda$  given  $\Lambda \neq 1$  is  $a/b$ , we set  $a = b$  to center  $\pi$  on  $\Lambda = 1$ , which appears to be appropriate for a fair comparison. With this choice, the conditional variance of  $\Lambda$  given  $\Lambda \neq 1$  is  $a/b^2 = 1/b$ , so  $b$  can be interpreted as a measure of prior conditional precision. Note that  $b = 1$  spreads  $\pi$  over the positive reals in such a way that  $\Lambda = 0$  lies exactly one standard deviation away from  $\Lambda = 1$ .

Figure 8 reports the posterior probability of criticality as a function of (the decimal logarithm of)  $b$ : it is apparent that the posterior probability of criticality is never larger than 50%; it gets very close, for extremely large values of  $b$ , but this is only because for such values  $\pi$  spikes on  $\Lambda = 1$  and this makes  $\Lambda \neq 1$  virtually indistinguishable from  $\Lambda = 1$ . For a very broad range of choices of  $b$  (about thirty orders of magnitude), the posterior probability of criticality is negligible. In particular, for the special choice  $b = 1$  discussed above, we get  $\Pr(\Lambda = 1 \mid V_{1:n} = v_{1:n}) \approx 4.32 \times 10^{-31}$ .

### Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

### Conflicts of Interest

The authors declare that they have no conflicts of interest.

### Acknowledgments

We are deeply indebted, for many reasons, to Alex Graudenzi, Alessandro Filisetti, and Chiara Damiani. We gratefully acknowledge the useful discussions on various aspects of the criticality hypothesis with Andrea Roli and on the size distribution of avalanches in random Boolean networks with Marina Luce Di Stefano.

### References

- [1] S. A. Kauffman, *The Origins of Order*, Oxford University Press, 1993.
- [2] S. A. Kauffman, *At Home in the Universe*, Oxford University Press, 1995.
- [3] S. Benedettini, M. Villani, A. Roli et al., “Dynamical regimes and learning properties of evolved Boolean networks,” *Neuro-computing*, vol. 99, pp. 111–123, 2013.
- [4] C. G. Langton, “Computation at the edge of chaos: phase transitions and emergent computation,” *Physica D: Nonlinear Phenomena*, vol. 42, no. 1–3, pp. 12–37, 1990.



- [5] C. G. Langton, "Life at the edge of chaos," in *Artificial Life II*, C. G. Langton, C. Taylor, J. D. Farmer, and S. Rasmussen, Eds., pp. 41–91, Addison-Wesley, 1992.
- [6] N. H. Packard, "Adaptation toward the edge of chaos," in *Dynamic Patterns in Complex Systems*, pp. 293–301, World Scientific, 1988.
- [7] Y. Bar-Yam, *Dynamics of Complex Systems*, Addison-Wesley, 1997.
- [8] G. Nicolis and C. Nicolis, *Foundations of Complex Systems: Nonlinear Dynamics, Statistical Physics, Information and Prediction*, World Scientific, 2007.
- [9] G. Nicolis and I. Prigogine, *Exploring Complexity: An Introduction*, W. H. Freeman, 1989.
- [10] J. J. Binney, N. J. Dowrick, A. J. Fisher, and M. E. J. Newman, *The Theory of Critical Phenomena*, Oxford University Press, 1992.
- [11] I. Shmulevich, S. A. Kauffman, and M. Aldana, "Eukaryotic cells are dynamically ordered or critical but not chaotic," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 102, no. 38, pp. 13439–13444, 2005.
- [12] C. Torres-Sosa, S. Huang, and M. Aldana, "Criticality is an emergent property of genetic networks that exhibit evolvability," *PLoS Computational Biology*, vol. 8, no. 9, article e1002669, 2012.
- [13] J. Hidalgo, J. Grilli, S. Suweis, M. A. Munoz, J. R. Banavar, and A. Maritan, "Information-based fitness and the emergence of criticality in living systems," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 111, no. 28, pp. 10095–10100, 2014.
- [14] M. Aldana, "Boolean dynamics of networks with scale-free topology," *Physica D: Nonlinear Phenomena*, vol. 185, no. 1, pp. 45–66, 2003.
- [15] F. Bailly and G. Longo, "Extended critical situations: the physical singularity of life phenomena," *Journal of Biological Systems*, vol. 16, no. 2, pp. 309–336, 2008.
- [16] T. Mora and W. Bialek, "Are biological systems poised at criticality?," *Journal of Statistical Physics*, vol. 144, no. 2, pp. 268–302, 2011.
- [17] A. Roli, M. Villani, A. Filisetti, and R. Serra, "Dynamical criticality: overview and open questions," *Journal of Systems Science and Complexity*, vol. 31, no. 3, pp. 647–663, 2018.
- [18] S. Kauffman, "Homeostasis and differentiation in random genetic control networks," *Nature*, vol. 224, no. 5215, pp. 177–178, 1969.
- [19] U. Bastolla and G. Parisi, "Relevant elements, magnetization and dynamical properties in Kauffman networks: a numerical study," *Physica D: Nonlinear Phenomena*, vol. 115, no. 3–4, pp. 203–218, 1998.
- [20] M. Aldana, S. Coppersmith, and L. P. Kadanoff, "Boolean dynamics with random couplings," in *Perspectives and Problems in Nonlinear Science*, E. Kaplan, J. E. Marsden, and K. R. Sreenivasan, Eds., pp. 23–89, Springer, 2003.
- [21] M. L. Di Stefano, M. Villani, L. La Rocca, S. A. Kauffman, and R. Serra, "Dynamically critical systems and power-law distributions: avalanches revisited," in *WIVACE 2015: Advances in Artificial Life, Evolutionary Computation and Systems Chemistry*, F. Rossi, F. Mavelli, P. Stano, and D. Caivano, Eds., vol. 587 of Communications in Computer and Information Science, pp. 29–39, Springer, 2016.
- [22] R. Serra, M. Villani, A. Graudenzi, and S. A. Kauffman, "Why a simple model of genetic regulatory networks describes the distribution of avalanches in gene expression data," *Journal of Theoretical Biology*, vol. 246, no. 3, pp. 449–460, 2007.
- [23] R. Serra, M. Villani, and A. Semeria, "Genetic network models and statistical properties of gene expression data in knock-out experiments," *Journal of Theoretical Biology*, vol. 227, no. 1, pp. 149–157, 2004.
- [24] T. R. Hughes, M. J. Marton, A. R. Jones et al., "Functional discovery via a compendium of expression profiles," *Cell*, vol. 102, no. 1, pp. 109–126, 2000.
- [25] P. Kemmeren, K. Sameith, L. A. L. van de Pasch et al., "Large-scale genetic perturbations reveal regulatory networks and an abundance of gene-specific repressors," *Cell*, vol. 157, no. 3, pp. 740–752, 2014.
- [26] B. Derrida and H. Flyvbjerg, "The random map model: a disordered model with deterministic dynamics," *Journal de Physique*, vol. 48, no. 6, pp. 971–978, 1987.
- [27] B. Derrida and Y. Pomeau, "Random networks of automata: a simple annealed approximation," *Europhysics Letters (EPL)*, vol. 1, no. 2, pp. 45–49, 1986.
- [28] P. Rämö, S. Kauffman, J. Kesseli, and O. Yli-Harja, "Measures for information propagation in Boolean networks," *Physica D: Nonlinear Phenomena*, vol. 227, no. 1, pp. 100–104, 2007.
- [29] D. Noble, *The Music of Life: Biology beyond the Genome*, Oxford University Press, 2006.
- [30] R. Serra and M. Villani, *Modelling protocells*, Springer, 2017.
- [31] K. Kaneko, *Life: An Introduction to Complex Systems Biology*, Springer, 2006.
- [32] A. Roli, S. Benedettini, R. Serra, and M. Villani, "Analysis of attractor distances in random Boolean networks," in *Neural Nets WIRN10*, B. Apolloni, S. Bassi, A. Esposito, and C. F. Morabito, Eds., pp. 201–208, IOS Press, 2011.
- [33] A. Roli, M. Villani, R. Caprari, and R. Serra, "Identifying critical states through the relevance index," *Entropy*, vol. 19, no. 2, p. 73, 2017.
- [34] T. I. Lee, N. J. Rinaldi, F. Robert et al., "Transcriptional regulatory networks in *Saccharomyces cerevisiae*," *Science*, vol. 298, no. 5594, pp. 799–804, 2002.
- [35] K. D. MacIsaac, T. Wang, D. B. Gordon, D. K. Gifford, G. D. Stormo, and E. Fraenkel, "An improved map of conserved regulatory sites for *Saccharomyces cerevisiae*," *BMC Bioinformatics*, vol. 7, no. 1, p. 113, 2006.
- [36] P. T. Monteiro, N. D. Mendes, M. C. Teixeira et al., "YEAS-TRACT-DISCOVERER: new tools to improve the analysis of transcriptional regulatory associations in *Saccharomyces cerevisiae*," *Nucleic Acids Research*, vol. 36, Supplement 1, pp. D132–D136, 2007.
- [37] D. Marbach, J. C. Costello, R. Küffner et al., "Wisdom of crowds for robust gene network inference," *Nature Methods*, vol. 9, no. 8, pp. 796–804, 2012.
- [38] G. Stolovitzky, D. Monroe, and A. Califano, "Dialogue on reverse-engineering assessment and methods: the DREAM of high-throughput pathway inference," *Annals of the New York Academy of Sciences*, vol. 1115, no. 1, pp. 1–22, 2007.
- [39] S. Ma, P. Kemmeren, D. Gresham, and A. Statnikov, "De-novo learning of genome-scale regulatory networks in *S. cerevisiae*," *PLoS One*, vol. 9, no. 9, article e106479, 2014.
- [40] A.-L. Barabási and R. Albert, "Emergence of scaling in random networks," *Science*, vol. 286, no. 5439, pp. 509–512, 1999.
- [41] M. E. J. Newman, "The structure and function of complex networks," *SIAM Review*, vol. 45, no. 2, pp. 167–256, 2003.

- [42] L. Glass and S. A. Kauffman, "The logical analysis of continuous, non-linear biochemical control networks," *Journal of Theoretical Biology*, vol. 39, no. 1, pp. 103–129, 1973.
- [43] A. Graudenzi, R. Serra, M. Villani, A. Colacci, and S. A. Kauffman, "Robustness analysis of a Boolean model of gene regulatory network with memory," *Journal of Computational Biology*, vol. 18, no. 4, pp. 559–577, 2011.
- [44] D. Sapienza, M. Villani, and R. Serra, "Dynamical properties of a gene-protein model," in *WTVACE 2017: Artificial Life and Evolutionary Computation*, M. Pelillo, I. Poli, A. Roli, R. Serra, D. Slanzi, and M. Villani, Eds., vol. 830 of Communications in Computer and Information Science, pp. 142–152, Springer, 2018.
- [45] B. Drossel, "Random Boolean networks," in *Reviews of Nonlinear Dynamics and Complexity*, H. G. Schuster, Ed., pp. 69–110, Wiley, 2008.
- [46] B. Luque and R. V. Solé, "Lyapunov exponents in random Boolean networks," *Physica A: Statistical Mechanics and its Applications*, vol. 284, no. 1–4, pp. 33–45, 2000.
- [47] M. Villani, D. Campioli, C. Damiani, A. Roli, A. Filisetti, and R. Serra, "Dynamical regimes in non-ergodic random Boolean networks," *Natural Computing*, vol. 16, no. 2, pp. 353–363, 2017.
- [48] M. Prokopenko, J. T. Lizier, O. Obst, and X. R. Wang, "Relating Fisher information to order parameters," *Physical Review E*, vol. 84, no. 4, article 041116, 2011.
- [49] S. E. Harris, B. K. Sawhill, A. Wuensche, and S. Kauffman, "A model of transcriptional regulatory networks based on biases in the observed regulation rules," *Complexity*, vol. 7, no. 4, 40 pages, 2002.
- [50] R. Van Der Hofstad and M. Keane, "An elementary proof of the hitting time theorem," *American Mathematical Monthly*, vol. 115, no. 8, pp. 753–756, 2008.
- [51] T. E. Harris, *The Theory of Branching Processes*, Springer, 1963.
- [52] R Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, 2018.
- [53] B. Efron, *The Jackknife, the Bootstrap and Other Resampling Plans*, SIAM, 1982.
- [54] R. G. Miller, "The jackknife-a review," *Biometrika*, vol. 61, no. 1, pp. 1–15, 1974.
- [55] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, Wiley, 2nd edition, 2006.
- [56] P. Bremaud, *Markov Chains: Gibbs Fields, Monte Carlo Simulation, and Queues*, Springer, 1999.
- [57] A. Graudenzi, R. Serra, M. Villani, C. Damiani, A. Colacci, and S. A. Kauffman, "Dynamical properties of a Boolean model of gene regulatory network with memory," *Journal of Computational Biology*, vol. 18, no. 10, pp. 1291–1303, 2011.
- [58] E. Balleza, E. R. Alvarez-Buylla, A. Chaos, S. Kauffman, I. Shmulevich, and M. Aldana, "Critical dynamics in genetic regulatory networks: examples from four kingdoms," *PLoS One*, vol. 3, no. 6, article e2456, 2008.
- [59] A. Wagner, "Estimating coarse gene network structure from large-scale gene perturbation data," *Genome Research*, vol. 12, no. 2, pp. 309–315, 2002.
- [60] R. Serra, M. Villani, A. Graudenzi, A. Colacci, and S. Kauffman, "The simulation of gene knock-out in scale-free random Boolean models of genetic networks," *Networks and Heterogeneous Media*, vol. 3, no. 2, pp. 333–343, 2008.
- [61] I. Vopálenská, M. Hulková, B. Janderová, and Z. Palková, "The morphology of *Saccharomyces cerevisiae* colonies is affected by cell adhesion and the budding pattern," *Research in Microbiology*, vol. 156, no. 9, pp. 921–931, 2005.
- [62] C. Damiani, A. Graudenzi, M. Villani, R. Serra, A. Colacci, and S. A. Kauffman, "On the fate of perturbations in critical random Boolean networks," *Proceedings of the European Conference on Complex Systems (ECCS09)*, 2009, pp. 1–10, University of Warwick, UK, 2009, CD-Rom paper 85.
- [63] C. Damiani, R. Serra, M. Villani, S. A. Kauffman, and A. Colacci, "Cell-cell interaction and diversity of emergent behaviours," *IET Systems Biology*, vol. 5, no. 2, pp. 137–144, 2011.
- [64] R. Serra, M. Villani, C. Damiani, A. Graudenzi, and A. Colacci, "The diffusion of perturbations in a model of coupled random Boolean networks," in *ACRI 2008: Cellular Automata*, H. Umeo, S. Morishita, K. Nishinari, T. Komatsuzaki, and S. Bandini, Eds., vol. 5191 of Lecture Notes in Computer Science, pp. 315–322, Springer, 2008.
- [65] R. Serra, M. Villani, C. Damiani, A. Graudenzi, A. Colacci, and S. A. Kauffman, "Interacting random Boolean networks," *Proceedings of ECCS07: European Conference on Complex Systems*, 2007, pp. 1–15, Technische Universität Dresden, Germany, 2007, CD-Rom paper #35.
- [66] M. Villani, R. Serra, P. Ingrami, and S. A. Kauffman, "Coupled random Boolean network forming an artificial tissue," in *ACRI 2006: Cellular Automata*, S. Yacoubi, B. Chopard, and S. Bandini, Eds., vol. 4173 of Lecture Notes in Computer Science, pp. 548–556, Springer, 2006.

